

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **06266399 A**(43) Date of publication of application: **22.09.94**

(51) Int. Cl.

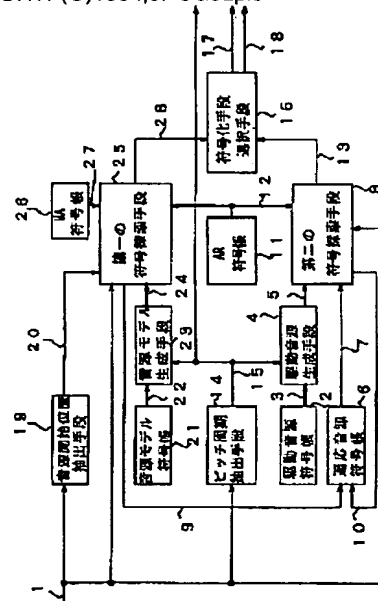
G10L 9/14**G10L 9/18**(21) Application number: **05049474**(22) Date of filing: **10.03.93**(71) Applicant: **MITSUBISHI ELECTRIC CORP**(72) Inventor: **SEZA KATSUSHI
TAZAKI HIROHISA****(54) ENCODING DEVICE AND SPEECH ENCODING
AND DECODING DEVICE****(57) Abstract:**

PURPOSE: To improve the quality of a decoded speech by the device which separates an input speech into a spectrum parameter and a sound source signal and performs encoding at every frame having fixed time length.

CONSTITUTION: The devices are equipped with a 1st code search means which searches for a combination of a sound source model code word with a spectrum code word minimizing the distortion of a synthesized speech and the input speech and outputs a 1st code search result 28, an adaptive sound source code book 6 which contains a 1st or 2nd quantized sound source signal 9 or 10 as an adaptive sound source signal, a driving sound source generating means 4 which generates a driving sound source signal from driving sound source code words in the driving sound source code book 2, a 2nd code search means which searches for a driving sound source code word minimizing the distortion of both the input speech and the synthesized speech generated by using a spectrum code word minimizing the spectrum distortion of a 2nd quantized sound source signal 10 generated by using an adaptive sound source signal and the driving sound source signal and outputs it as a 2nd

code search result 13, and an encoding means selecting means 16 which outputs one of the 1st and 2nd code search results.

COPYRIGHT: (C)1994,JPO&Japio



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平6-266399

(43) 公開日 平成 6 年 (1994) 9 月 22 日

(51) Int. Cl. ⁵

G10L 9/14

9/18

識別記号

庁内整理番号

G 8946-5H

J 8946-5H

E 8946-5H

F I

技術表示箇所

審査請求 未請求 請求項の数 7 O L (全20頁)

(21) 出願番号 特願平5-49474

(22) 出願日 平成 5 年 (1993) 3 月 10 日

(71) 出願人 000006013

三菱電機株式会社

東京都千代田区丸の内二丁目 2 番 3 号

(72) 発明者 瀬座 勝志

鎌倉市大船五丁目 1 番 1 号 三菱電機株式

会社情報電子研究所内

(72) 発明者 田崎 裕久

鎌倉市大船五丁目 1 番 1 号 三菱電機株式

会社情報電子研究所内

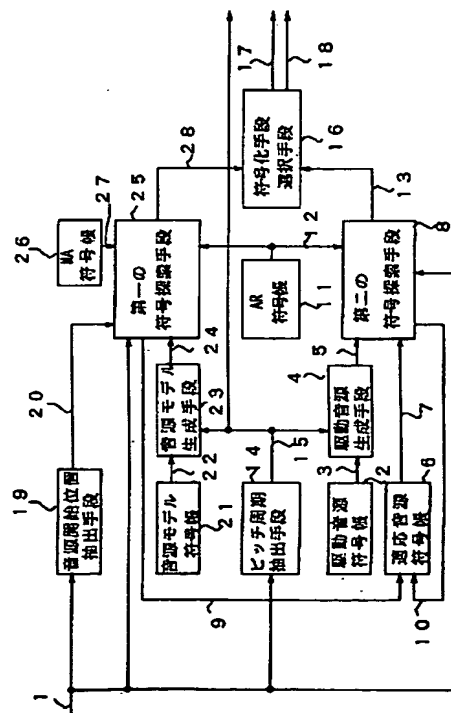
(74) 代理人 弁理士 高田 守

(54) 【発明の名称】 符号化装置及び音声符号化復号化装置

(57) 【要約】

【目的】 入力音声をスペクトルパラメータと音源信号に分離して固定時間長のフレーム毎に符号化する装置において、復号音声の品質を改善する。

【構成】 合成音声と入力音声の歪を最小にする音源モデル符号語とスペクトル符号語の組み合わせを探索し、第一の符号探索結果 28 を出力する第一の符号探索手段と、第一または第二の量子化音源信号 9、10 を適応音源信号として格納した適応音源符号帳 6 と、駆動音源符号帳 2 内の駆動音源符号語より駆動音源信号を生成する駆動音源生成手段 4 と、適応音源信号と駆動音源信号を用いて生成した第二の量子化音源信号 10 とのスペクトル歪を最小にするスペクトル符号語を用いて生成した合成音声と入力音声の歪を最小にする駆動音源符号語を探索し、第二の符号探索結果 13 として出力する第二の符号探索手段と、第一と第二の符号探索結果のうち一方を出力する符号化手段選択手段 16 を備える。



【特許請求の範囲】

【請求項 1】 入力信号と合成信号の歪を最小にするようなモデルを探索してその探索結果を用いて入力信号を符号化する符号化装置において、

入力信号よりピッチ周期を抽出するピッチ周期抽出手段と、

前記ピッチ周期抽出手段により抽出されたピッチ周期に対して複数のモデルの中から入力信号に適合するひとつのモデルを選択して入力信号を符号化する第 1 の符号化手段と、

前記第 1 の符号化手段とは異なる方式により入力信号を符号化する第 2 の符号化手段と、

前記第 1 の符号化手段と第 2 の符号化手段による符号化結果を比較し、その比較結果に基づいて第 1 の符号化手段と第 2 の符号化手段のいずれの符号化結果を選択して出力する選択手段とを有することを特徴とする符号化装置。

【請求項 2】 入力音声、声道のスペクトルパラメータと音源信号に分離して固定時間長のフレーム毎に符号化復号化する音声符号化復号化装置において、

符号化部に、入力音声よりピッチ周期を抽出し音源モデル生成手段と駆動音源生成手段に出力するピッチ周期抽出手段と、典型的なスペクトルパラメータをスペクトル符号語として複数個格納したスペクトル符号帳と、一ピッチ周期の音源信号を表す音源モデルのパラメータの典型的なものを音源モデル符号語として複数個格納した音源モデル符号帳と、前記音源モデル符号帳内の音源モデル符号語から生成した一ピッチ周期の音源信号を前記ピッチ周期で繰り返したものを第一の量子化音源信号として第一の符号探索手段に出力する音源モデル生成手段

と、前記第一の量子化音源信号と前記スペクトル符号帳内のスペクトル符号語から生成した合成音声と入力音声の歪を最小にする音源モデル符号語とスペクトル符号語の組み合わせを探索し、その探索結果を第一の符号探索結果としてその時の歪と共に符号化手段選択手段に出力し、この組み合わせにおける第一の量子化音源信号を適応音源符号帳に出力する第一の符号探索手段と、

先行フレームにおいて第一の符号探索手段が出力した第一の量子化音源信号あるいは第二の符号探索手段が出力した第二の量子化音源信号を適応音源信号として格納した前記適応音源符号帳と、予め用意された信号を駆動音源符号語として複数個格納した駆動音源符号帳と、前記駆動音源符号帳内の駆動音源符号語を前記ピッチ周期で繰り返した駆動音源信号を生成する駆動音源生成手段

と、前記適応音源信号と前記駆動音源信号から第二の量子化音源信号を生成し、前記スペクトル符号帳より当該フレームにおけるスペクトル歪を最小にするスペクトル符号語を選択し、このスペクトル符号語と前記第二の量子化音源信号を用いて合成音声を生

成し、合成音声と入力音声の歪を最小にする適応音源信号を探索結果を第二の符号探索結果としてその時の歪と共に符号化手段選択手段に出力し、第二の量子化音源信号を前記適応音源符号帳に出力する第二の符号探索手段と、前記第一の符号探索結果と前記第二の符号探索結果の内、より小さい歪を持つ方を当該フレームの符号化結果として選択し出力するとともに、どちらの符号探索結果を選択したのかを符号化手段選択信号として出力する符号化手段選択手段を備え、

復号化部に、符号化部と同じスペクトル符号帳と、符号化部と同じ音源モデル符号帳と、

符号化部より入力された符号化結果に対応する前記音源モデル符号帳内の音源モデル符号語と符号化部より入力されたピッチ周期から第一の量子化音源信号を生成して出力する符号化部と同じ音源モデル生成手段と、符号化部と同じ適応音源符号帳と、符号化部と同じ駆動音源符号帳と、前記第一の量子化音源信号と符号化部より入力された符号化結果に対応する前記スペクトル符号帳内のスペクトル符号語を用いて復号音声を生

成し、前記適応音源符号帳に前記第一の量子化音源信号を出力する第一の復号化手段と、前記符号化結果に対応する前記駆動音源符号帳内の駆動音源符号語と前記ピッチ周期から駆動音源信号を生成する駆動音源生成手段と、前記適応音源符号帳内の適応音源信号と前記駆動音源生成手段の出力する駆動音源信号より生成した第二の量子化音源信号と前記スペクトル符号語より復号音声を生

成し、第二の量子化音源信号を前記適応音源符号帳に出力する第二の復号化手段と符号化部より入力された符号化手段選択信号に従い第一の復号化手段の復号音声と第二の復号化手段の復号音声を選択する復号化手段選択手段とを備えることを特徴とする音声符号化復号化装置。

【請求項 3】 上記音声符号化復号化装置において、符号化部と復号化部の適応音源符号帳内に、第一の符号探索手段が選択された場合の第一の量子化音源信号を格納する第一の音源記憶手段と、第二の符号探索手段が選択された場合の第二の量子化音源信号を格納する第二の音源記憶手段と、適応音源信号として第一の量子化音源信号と第二の量子化音源信号を切り換えて出力する切換手段を備え、符号化部に、前記適応音源信号と駆動音源信号から第二の量子化音源信号を生成し、この第二の量子化音源信号と A R 符号語を用いて合成音声を生

成し、この合成音声と入力音声の歪を小さくする適応音源信号を選択し、その選択結果を第二の符号探索結果に含めて出力するようにした第二の符号探索手段を備え、復号化部に、符号化部より入力された符号化結果に従い適応音源符号帳内の第一の量子化音源信号または第二の量子化音源信号を適応音源信号として選択して用いるようにした第二の復号化手段を備えることを特徴とする請求項 2 記載の音声符号化復号化装置。

【請求項 4】 入力音声、声道のスペクトルパラメータと音源信号に分離して固定時間長のフレーム毎に符号

化復号化する音声符号化復号化装置において、符号化部に、少なくとも、典型的なスペクトルパラメータをスペクトル符号語として複数個格納したスペクトル符号帳と、一ピッチ周期の音源信号を表す音源モデルのパラメータの典型的なものを音源モデル符号語として複数個格納した音源モデル符号帳と、予め用意された信号を駆動音源符号語として複数個格納した駆動音源符号帳とのいずれかを備え、少なくとも符号化部に備えたスペクトル符号帳または音源モデル符号帳または駆動音源符号帳のいずれかに、ピッチ周期に対応する複数個の副符号帳と、ピッチ周期抽出手段より入力されたピッチ周期に応じて符号化復号化処理に用いる前記副符号帳を切り換える副符号帳切換手段を備え、復号化部に、復号化部より入力されたピッチ周期に応じて副符号帳を切り換える符号化部と同一のスペクトル符号帳または音源モデル符号帳または駆動音源符号帳を備えることを特徴とする音声符号化復号化装置。

【請求項 5】 入力音声、スペクトルパラメータと音源信号に分離して固定時間長のフレーム毎に符号化する音声符号化復号化装置において、符号化部に、入力音声よりピッチ周期を抽出し音源モデル生成手段と駆動音源生成手段とスペクトル符号帳と音源モデル符号帳に出力するピッチ周期抽出手段と、典型的なスペクトルパラメータをスペクトル符号語として複数個格納したスペクトル符号帳と、一ピッチ周期の音源信号を表す音源モデルのパラメータの典型的なものを音源モデル符号語として複数個格納した音源モデル符号帳と、前記スペクトル符号帳または前記音源モデル符号帳に、ピッチ周期に対応する複数個の副符号帳と、ピッチ周期抽出手段より入力されたピッチ周期に応じて符号化復号化処理に用いる前記副符号帳を切り換える副符号帳切換手段を備え、前記音源モデル符号帳内の音源モデル符号語から生成した一ピッチ周期の音源信号を前記ピッチ周期で繰り返したものを量子化音源信号として符号探索手段に出力する音源モデル生成手段と、前記量子化音源信号と前記スペクトル符号帳内のスペクトル符号語から生成した合成音声と入力音声の歪を最小にする音源モデル符号語とスペクトル符号語の組み合わせを探索し、その探索結果を符号探索結果として出力する符号探索手段と、復号化部に、符号化部と同じスペクトル符号帳と、符号化部と同じ音源モデル符号帳と、符号化部より入力された符号化結果に対応する前記音源モデル符号帳内の音源モデル符号語と符号化部より入力されたピッチ周期から量子化音源信号を生成して復号化手段に出力する符号化部と同じ音源モデル生成手段と、前記量子化音源信号と符号化部より入力された符号化結果に対応する前記スペクトル符号帳内のスペクトル符号語を用いて復号音声を生成し、前記適応音源符号帳に前記量子化音源信号を出力する復号化手段を備えることを特徴とする音声符号化

復号化装置。

【請求項 6】 上記音声符号化復号化装置において、符号化部に、当該フレームが無声フレームから初めて有声フレームに変わったフレームの場合、有声音が開始する位置を音源開始位置として入力音声より抽出し、当該フレームに先行する無声フレームにおいて復号化部に出力する音源開始位置抽出手段と、復号化部に、第一の量子化音源信号を符号化部より入力された音源開始位置に同期して復号音声を生成するようにした第一の復号化手段を備えることを特徴とする請求項 2 記載の音声符号化復号化装置。

【請求項 7】 入力音声、スペクトルパラメータと音源信号に分離して固定時間長のフレーム毎に符号化する音声符号化復号化装置において、符号化部に、入力音声よりピッチ周期を抽出し音源モデル生成手段と駆動音源生成手段に出力するピッチ周期抽出手段と、典型的なスペクトルパラメータをスペクトル符号語として複数個格納したスペクトル符号帳と、一ピッチ周期の音源信号を表す音源モデルのパラメータの典型的なものを音源モデル符号語として複数個格納した音源モデル符号帳と、前記音源モデル符号帳内の音源モデル符号語から生成した一ピッチ周期の音源信号を前記ピッチ周期で繰り返したものを量子化音源信号として符号探索手段に出力する音源モデル生成手段と、前記量子化音源信号と前記スペクトル符号帳内のスペクトル符号語から生成した合成音声と入力音声の歪を最小にする音源モデル符号語とスペクトル符号語の組み合わせを探索し、その探索結果を符号探索結果として出力する符号探索手段と、当該フレームが無声フレームから初めて有声フレームに変わったフレームの場合、有声音が開始する位置を音源開始位置として入力音声より抽出し、当該フレームに先行する無声フレームにおいて復号化部に出力する音源開始位置抽出手段と、復号化部に、符号化部と同じスペクトル符号帳と、符号化部と同じ音源モデル符号帳と、符号化部より入力された符号化結果に対応する前記音源モデル符号帳内の音源モデル符号語と符号化部より入力されたピッチ周期から量子化音源信号を生成して復号化手段に出力する符号化部と同じ音源モデル生成手段と、当該フレームが無声フレームから初めて有声フレームに変わったフレームの場合、符号化部より入力された音源開始位置に同期した前記量子化音源信号と符号化部より入力された符号化結果に対応する前記スペクトル符号帳内のスペクトル符号語を用いて復号音声を生成し、前記適応音源符号帳に前記量子化音源信号を出力する復号化手段を備えることを特徴とする音声符号化復号化装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 この発明は、音声をデジタル伝送あるいは蓄積する場合に用いられる音声符号化復号化

装置に関するものである。

【 0 0 0 2 】

【従来の技術】入力音声、スペクトルパラメータと音源信号に分離して固定時間長のフレーム毎に符号化する従来の音声符号化復号化装置は、文献 1” 声門音源波モデルを用いた音声の分析合成方式の検討”（瀬座勝志、田崎裕久、中島邦男、日本音響学会秋季研究発表会、1 - 6 - 1 0、P P 2 0 9 - 2 1 0、1 9 9 1）により報告されている。この従来法においては、音源信号の符号化に声門音源波の微分波形上で定義される音源モデルを用い、スペクトルパラメータとして自己回帰係数（以下 A R と略す）及び移動平均係数（以下 M A と略す）を用いている。前述した文献では、2 ~ 3 K b p s 程度の低ビットレート音声伝送において高品質な復号音声を得る方式として声帯音源波モデルを用いた分析合成方式（F V Q - G A R M A）を検討している。そこでは、声帯音源波モデル、A R パラメータ、及び M A パラメータを全てベクトル量子化することにより、自然性の高い復号音声を得られることを明かにしている。しかし、声帯音源波モデルの予備選択に用いる音源ピーク位置の抽出誤りや、声帯音源波モデルをマッチングする位置（音源位置）の伝送を行わないことにより復号音声に劣化を生ずる場合があった。従来の F V Q - G A R M A 方式は、声帯音源波モデルで A R M A フィルタを駆動することによって有声音を生成するものである。声帯音源波モデルには声帯音源波の微分波形上で定義されるモデルを用いている。符号化部では声帯音源波モデル、A R 及び M A パラメータをベクトル量子化する際に、予備選択された各コードの全ての組み合わせの中から S N R s e g が最大になる組み合わせをフレームに一組選択する。復号化部では各フレームで得られたコードをそれぞれ補間しながら合成する。

【 0 0 0 3 】図 1 4 及び図 1 5 は従来の分析合成方式（F V Q - G A R M A）による音声符号化復号化装置の構成図である。図 1 4 は符号化部を示し、図 1 5 は復号化部を示している。図 1 4 及び図 1 5 において、1 は入力音声、1 1 は A R 符号帳、1 2 は A R 符号語、1 4 はピッチ周期抽出手段、1 5 はピッチ周期、1 9 は音源開始位置抽出手段、2 0 は音源開始位置、2 1 は音源モデル符号帳、2 2 は音源モデル符号語、2 3 は音源モデル生成手段、2 6 は M A 符号帳、2 7 は M A 符号語、2 9 は音源モデル符号帳、3 0 は音源モデル符号語、3 1 は音源モデル生成手段、3 4 は M A 符号帳、3 5 は M A 符号語、3 7 は復号音声、4 4 は A R 符号帳、4 5 は A R 符号語、5 6 は符号化結果、5 7 は符号探索手段、5 8 は復号化手段、5 9 は量子化音源信号、6 0 は量子化音源信号である。

【 0 0 0 4 】まず、図 1 4 の符号化部について説明する。A R 符号帳 1 1 には典型的な A R を A R 符号語として複数個格納し、M A 符号帳 2 6 には典型的な M A を M

A 符号語として複数個格納し、音源モデル符号帳 2 1 には一ピッチ周期の音源信号を表す音源モデルのパラメータの典型的なものを音源モデル符号語として複数個格納してある。ピッチ周期抽出手段 1 4 は入力音声 1 よりピッチ周期 1 5 を抽出し出力する。音源開始位置抽出手段 1 9 は先行フレームが無声で当該フレームが有聲の場合、入力音声 1 より音源開始位置 2 0 を抽出し出力する。音源モデル生成手段 2 3 は音源モデル符号帳 2 1 より出力される音源モデル符号語 2 2 より生成される一ピッチ周期の音源信号をピッチ周期 1 5 で繰り返した信号を生成し、量子化音源信号 5 9 として出力する。符号探索手段 5 7 は、音源開始位置 2 0 と A R 符号語 1 2 と量子化音源信号 5 9 と M A 符号語 2 7 を用いて合成音声を生成し、入力音声 1 と合成音声の歪を最小にする A R 符号語と M A 符号語と音源符号語の組み合わせを探索し、符号化結果 5 6 として出力する。

【 0 0 0 5 】図 1 6 は有声音の先頭フレームでの符号探索手段 5 7 の動作を説明するものである。図において実線で示される量子化音源信号及び合成音声は当該フレームでの信号を、点線で示される量子化音源信号及び合成音声は次フレームでの信号を示す。符号探索手段 5 7 は、有聲フレームにおいてピッチ周期を単位とした入力音声を符号化する。当該フレームが有声音の先頭である場合は、音源開始位置 2 0 からピッチ周期単位で量子化音源信号 5 9 を並べた場合に当該フレームを超える範囲の入力音声 1 を当該フレームの符号化対象とし、合成音声を生成する。

【 0 0 0 6 】図 1 6 においては、音源開始位置 2 0 からピッチ周期 P 1、P 2 までが現在のフレームの符号化対象として合成音声生成される。符号化はピッチ周期を単位として行われるため、フレームの区切りとピッチ周期の区切りは一致せず、この例では、時刻 T 3 は時刻 F 2 とは一致せず、時刻 T 3 - 時刻 F 2 の時間だけ、現在のフレームを超えて符号化が行われる。有声音の先頭フレーム以外の有聲フレームの場合、量子化音源信号 5 9 を先行フレームでの量子化音源信号に引き続いて当該フレームを超えるまで並べて合成音声を生成する。図においては、P 2 というピッチ周期の次から次フレームのための音声合成が生成される。すなわち、P 3、P 4、P 5 というピッチ周期を用いて次フレームの符号化が行われ合成音声生成される。

【 0 0 0 7 】次に図 1 5 の復号化部について説明する。図において A R 符号帳 4 4、音源モデル符号帳 3 1、M A 符号帳 3 4 は、それぞれ符号化部における A R 符号帳 1 1、音源モデル符号帳 2 1、M A 符号帳 2 6 と同じものである。音源モデル生成手段 3 1 は、ピッチ周期 1 5 と符号化結果 5 6 に対応する音源モデル符号帳 2 9 内の音源モデル符号語 3 0 を用いて量子化音源信号 6 0 を生成する。復号化手段 5 8 は、量子化音源信号 6 0 と符号化結果 5 6 に対応する A R 符号帳 4 4 内の A R 符号語 4

5とMA符号帳34内のMA符号語35を用いて復号音声37を生成する。

【0008】図17は復号化手段58の動作を説明するものである。復号化手段58は量子化音源信号60を当該フレームの先頭から当該フレームを超えるまで並べて復号音声37を生成する。図においては、当該フレームの先頭からフレーム周期91、92、93を用いて復号音声37を生成する。復号音声の生成もピッチ周期単位で行われるため、フレームを超えて復号音声37が生成される場合がある。図17においては、ピッチ周期93の終了時刻S3はフレームの時刻F2を超えており、当該フレームの時刻F2が終了しても当該フレームのための復号音声37が時刻S3まで生成される。

【0009】次フレームでは、これまで復号された復号音声37に引き続いて点線で示される量子化音源信号60を並べて復号音声37を生成する。図17においては、ピッチ周期94、95、96が次フレームの復号音声生成のために用いられる。この次フレームの復号音声生成もピッチ周期単位で行われるため、前のフレームの復号音声生成が前のフレームを超えて行われる場合には、図17次に示すように次フレームにおいても、ずれたまま復号音声37を生成する。

【0010】

【発明が解決しようとする課題】従来の音声符号化復号化装置は、声帯音源波コードの予備選択を音源ピーク位置と過去のフレームで選択された声帯音源波コードを基準として行っているが、語頭部分や過度部での音源ピーク位置の自動抽出には誤りが多く、予備選択がうまく働かない場合があった。図18(a)に残差波形、図18(b)(c)に声帯音源波モデルの微分波形を示す。音源ピーク位置が正しく抽出されている場合(b)に比べ、誤って抽出された場合(c)は声帯音源波コードの選択を誤り、SNRsegは急速に劣化する。このように、音源信号の符号化に音源モデルを用いた量子化音源信号のみを使っているために音源モデルの適合の悪い話者の場合に復号音声の品質が劣化する場合があった。また、ピッチ周期に応じて様態が異なるMAと音源モデルに対して各々固定の符号帳を用いて量子化するために復号音声の品質が劣化する場合があった。また、符号化部ではピッチ長を補間により微調整しながら音源を繰り返した場合にSNRsegが最大になるように有声音の先頭の音源位置とそのピッチ長を決定しているが、この音源位置を復号化部に伝送しない構成のため、符号化部と復号化部で各コードを補間した結果に大きな差異を生じ復号音声品質が劣化する例があった。すなわち、復号化部に有声音の先頭フレームにおける音源開始位置が伝送されないために、図16及び図17に示すように、符号化部と復号化部において同一フレーム内に含まれる音源モデルの数が異なる場合がある。この様なフレームにおいてパワーやピッチ周期の変動が大きいと、復号化部の

第一の量子化音源信号は符号化部の第一の量子化音源信号との間に大きな差異を生じ、復号音声の品質が劣化する場合があった。

【0011】本発明は上記課題を解消するためになされたもので、復号音声の品質を向上させることを目的としている。

【0012】

【課題を解決するための手段】この発明の請求項1記載の発明に係る音声符号化復号化装置は、音質劣化を低減するため、たとえばGARMAとCELPとのマルチモード化を行い、CELP系で用いられている適応コードと駆動音源コードを用いた音源(図18(d)参照)を用いてSNRsegを確保し、FVQ-GARMAを用いた場合と適応コードブックと駆動音源コードブックを用いた場合でSNRsegの良い方を選択するようにしたものである。

【0013】この発明の請求項2記載の発明に係る音声符号化復号化装置は、符号化部に、第一の量子化音源信号とスペクトル符号帳内のスペクトル符号語から生成した合成音声と入力音声の歪を最小にする音源モデル符号語とスペクトル符号語の組み合わせを探索し、その探索結果を第一の符号探索結果としてその時の歪と共に符号化手段選択手段に出力し、第一の量子化音源信号を適応音源符号帳に出力する第一の符号探索手段と、先行フレームにおいて求めた第一の量子化音源信号または第二の量子化音源信号を適応音源信号として格納した適応音源符号帳と、予め用意された音源信号を駆動音源符号語として複数個格納した駆動音源符号帳と、前記駆動音源符号帳内の駆動音源符号語を前記ピッチ周期で繰り返した駆動音源信号を生成する駆動音源生成手段と、前記適応音源符号帳内の適応音源信号と前記駆動音源信号より生成される第二の量子化音源信号と前記スペクトル符号帳内のスペクトル符号語を用いて生成した合成音声と入力音声の歪を最小にする駆動音源符号語を探索し、その探索結果を第二の符号探索結果としてその時の歪と共に符号化手段選択手段に出力し、第二の量子化音源信号を適応音源符号帳に出力する第二の符号探索手段と、前記第一の符号探索結果と前記第二の符号探索結果の内、より小さい歪を持つ方を符号化結果として選択し、当該フレームの符号化結果として出力するとともに、どちらの符号探索結果を選択したのかを符号化手段選択信号として出力する符号化手段選択手段を備え、復号化部に、符号化部より入力された符号化手段選択信号に従い第一の復号化手段と第二の復号化手段を選択する復号化手段選択手段と、符号化部と同じスペクトル符号帳と、符号化部と同じ音源モデル符号帳と、第一の量子化音源信号と符号化部より入力された符号化結果に対応する前記スペクトル符号帳内のスペクトル符号語を用いて復号音声37を生成する第一の復号化手段と、符号化部と同じ適応音源符号帳と、符号化部と同じ駆動音源符号帳と、前記符号化

結果に対応する前記駆動音源符号帳内の駆動音源符号語と前記ピッチ周期から駆動音源信号を生成する駆動音源生成手段と、前記スペクトル符号語と前記適応音源符号帳内の適応音源信号と前記駆動音源生成手段の出力する駆動音源信号より復号音声の生成する第二の復号化手段を備える。

【0014】本発明の請求項3記載の発明に係わる音声符号化復号化装置は、符号化部と復号化部の適応音源符号帳に、第一の量子化音源信号を保持する第一の音源記憶手段と第二の量子化音源信号を保持する第二の音源記憶手段と、前記第一の音源記憶手段と前記第二の音源記憶手段を切り換える符号語切換手段を備え、合成音声と入力音声の歪を最小にする適応音源信号を前記適応音源符号帳より選択し、どちらを選択したかを第二の符号化結果に含めて出力する第二の符号探索手段を備える。

【0015】本発明の請求項4及び請求項5記載の発明に係わる音声符号化復号化装置の符号化部と復号化部のスペクトル符号帳、音源モデル符号帳及び駆動音源符号帳は、それぞれピッチ周期に対応する複数の副符号帳と、入力されるピッチ周期に応じて前記副符号帳を切り換える副符号帳切換手段を備える。

【0016】本発明の請求項6及び請求項7記載の発明に係わる音声符号化復号化装置は、有声音の先頭フレームの音源位置だけを、その直前の無声フレームにおいて伝送する構成とした。後続するフレームでは符号化部と復号化部においてピッチ長を補間しながら音源を繰り返すという同一の処理を行うため、フレーム毎に音源位置を伝送する必要はない。すなわち、符号化部に有声音の先頭フレームの場合、入力音声より音源開始位置を抽出し、先行する無声フレームにおいて復号化部に出力する音源開始位置抽出手段と、復号化部に、符号化部より入力された音源開始位置に第一の量子化音源信号を同期して復号音声の生成する第一の復号化手段を備える。

【0017】

【作用】請求項1記載の発明においては、たとえば、GARMAとCELPという異なる方式を用いて符号化しよりよい結果をもたらす方式を選択するので、符号化品質がどちらか一方の方式による場合よりも向上する。

【0018】また請求項2記載の発明においては、符号化手段選択手段は第一の符号探索手段と第二の符号探索手段のうち入力音声と合成音声の歪を小さくする方を選択する。

【0019】また、請求項3記載の発明においては、第二の符号探索手段は適応音源符号帳に保持されている第一の量子化音源信号と第二の量子化音源信号のうち入力音声と合成音声の歪を小さくする信号を適応音源信号として選択する。

【0020】また、請求項4、5記載の発明においては、スペクトル符号帳、音源モデル符号帳、駆動音源符号帳はピッチ周期に応じてそれぞれが持っている副符号

帳を切り換える。

【0021】また、請求項6、7記載の発明においては、有声音の先頭のフレームでの音源開始位置を復号化部に伝送する場合に有聲フレームに先行する無声フレームで伝送する。

【0022】

【実施例】

実施例1. 図1と図2はこの発明に係わる音声符号化復号化装置の一実施例の構成図であり、以下、本発明の動作をこの図において説明する。なお図1は符号化部を示し、図2は復号化部を示しており、図1と図2において図8と図9と同一の部分については同一符号を付し、説明を省略する。図において、2は駆動音源符号帳、3は駆動音源符号語、4は駆動音源生成手段、5は駆動音源信号、6は適応音源符号帳、7は適応音源信号、8は第一の符号探索手段、9は第一の量子化音源信号、10は第二の量子化音源信号、13は第二の符号探索結果、16は復号化手段選択手段、17は符号化手段選択信号、18は符号化結果、25は第二の符号探索手段、28は第一の符号探索結果、33は第一の復号化手段、38は符号化手段選択手段、39は符号化結果、40は適応音源符号帳、41は適応音源信号、42は第二の量子化音源信号、43は第二の復号化手段、46は駆動音源符号帳、47は駆動音源符号語、48は駆動音源生成手段、49は駆動音源信号である。

【0023】まず、符号化部について説明する。第一の符号探索手段25は、図10に示すように有声音の先頭フレームの場合は音源開始位置20に第一の量子化音源信号24を同期させ、この第一の量子化音源信号とAR符号語12とMA符号語27を用いて合成音声の生成し、それ以外の有聲フレームでは、先行フレームで得られた合成音声に引き続いて第一の量子化音源信号24を並べて合成音声の生成し、この合成音声と入力音声1の歪を最小にするAR符号語12とMA符号語27と音源モデル符号語22の組み合わせを探索し、その探索結果を符号化結果28とし、その歪と共に符号化手段選択手段16に出力し、またその組み合わせにおける第一の量子化音源信号24を第一の量子化音源信号9として適応音源符号帳6に出力する。

【0024】駆動音源符号帳2は例えばLPC残差信号の中で典型的な一ピッチ周期の信号やガウス性雑音信号を駆動音源符号語として複数個格納しておく。駆動音源生成手段4は駆動音源符号帳2内の駆動音源符号語3をピッチ周期15で繰り返した駆動音源信号5を生成する。以下この第一の符号探索手段25が行う動作方式をGARMA方式と呼ぶことにする。

【0025】第二の符号探索手段8は駆動音源信号5と適応音源符号帳6内の適応音源信号7より生成される量子化音源信号とスペクトル符号帳11内のスペクトル符号語12を用いて合成音声の生成し、合成音声と入力音

声 1 の歪を最小にする駆動音源符号語 3 とスペクトル符号語 1 2 の組み合わせを探索し、その探索結果を第二の符号探索結果 1 3 とし、その歪と共に符号化手段選択手段 1 6 に出力し、この組み合わせにおける第二の量子化音源信号 1 0 を適応音源符号帳 6 に出力する。以下この第二の符号探索手段 8 が符号化に用いる方式を CELP 方式と呼ぶことにする。

【0026】符号化手段選択手段 1 6 は第一の符号探索結果 2 8 と第二の符号探索結果 1 3 の内で歪の小さい方を当該フレームにおける符号化結果 1 8 として選択し、
10 どちらを選択したかを表す符号化手段選択信号 1 7 と前記符号化結果 1 8 を出力する。

【0027】図 3 は適応音源符号帳 6 の動作を説明した図である。適応音源符号帳 6 は当該フレームにおいて第一の符号探索手段が選択された場合は第一の量子化音源信号 9 を音源記憶手段 5 0 に格納し、第二の符号探索手段が選択された場合は第二の量子化音源信号 1 0 を音源記憶手段 5 0 に格納し、適応音源信号 7 として出力する。

【0028】従来の適応音源符号帳 6 は符号探索手段により、量子化音源信号を記憶しているのに対して、この実施例においては、適応音源符号帳 6 に記憶する量子化音源信号を第一の符号探索手段 2 5 により、出力された第一の量子化音源信号 9 と第二の符号探索手段 8 により出力された第二の量子化音源信号 1 0 のうちから最新の量子化音源信号を選択し音源記憶手段 5 0 に格納する。どちらが最新の量子化音源信号であるかは符号化手段選択手段から出力された符号化手段選択信号 1 7 により判定することができる。従って、適応音源符号帳 6 は符号化手段選択信号 1 7 を入力してスイッチを切り換えることにより、第一の量子化音源信号 9 と第二の量子化音源信号を 1 0 を切り換えて音源記憶手段 5 0 に入力する。こうして適応音源符号帳 6 は最新に用いられた量子化音源信号を第二の符号探索手段に供給することが可能になる。

【0029】次に図 2 の復号化部 1 b について説明する。図において駆動音源符号帳 4 6、適応音源符号帳 4 0 は、それぞれ符号化部 1 a における駆動音源符号帳 2、適応音源符号帳 6 と同一のものである。復号化手段選択手段 3 8 は符号化手段選択信号 1 7 に従い第一の復号化手段 3 3 と第二の復号化手段 4 3 のどちらかに、
40 符号化結果 1 8 をそのまま符号化結果 3 9 として出力する。

【0030】第一の復号化手段 3 3 は第一の量子化音源信号 3 2 と符号化結果 3 9 に対応する MA 符号帳 3 4 内の MA 符号語 3 5 とスペクトル符号帳 4 4 内の AR 符号語 4 5 を用いて復号音声 3 7 を生成し、第一の量子化音源信号 3 2 をそのまま第一の量子化音源信号 3 6 として適応音源符号帳 4 0 に出力する。

【0031】駆動音源生成手段 4 8 は、ピッチ周期 1 5
50

と符号化結果 3 9 に対応する駆動音源符号帳 4 6 内の駆動音源符号語 4 7 より駆動音源信号 4 9 を生成する。第二の復号化手段 4 3 は、適応音源符号帳 4 0 内の適応音源信号 4 1 と駆動音源信号 4 9 より生成される量子化音源信号と符号化結果 3 9 に対応するスペクトル符号帳 4 4 内の AR 符号語 4 5 を用いて復号音声 3 7 を生成し、前記量子化音源信号を適応音源符号帳に第二の量子化音源信号 4 2 として出力する。

【0032】以上のようにこの実施例では、第一の符号化探索手段は音源モデル符号帳 2 1 に格納された音源モデルを用いた量子化音源信号のみを用いて符号化を行っている。一方第二の符号探索手段は適応音源符号帳を用いることにより、直前に符号化した量子化音源信号との差分を用いることにより符号化を行っている。このようにこの符号化の方式が異なるふたつの符号探索手段をそれぞれ動作させ、その動作結果を比較することにより、符号化の歪の小さい方を選択する点がこの実施例の特徴である。すなわち、この実施例は GARMA 方式と CELP 方式の二つの方式を用い、二つの方式から得られた符号化結果のよりよい方を選択して出力することを特徴とするものである。

【0033】実施例 2. 上記実施例 1 においては、GARMA 方式と CELP 方式の二つの方式を用いて比較選択する場合を示したが、二つの符号化方式はこれらの方式に限るものではなく、その他の方式を利用するものでもかまない。あるいは同一方式のものであっても、一方に改良を加えたものや変更を加えたものであってもかまわない。更に、二つの方式の組み合わせに限らず、三つ以上の方式の組み合わせであってもかまわない。

【0034】実施例 3. 図 4 はこの発明に係わる音声符号化復号化装置の一実施例における適応音源符号帳 6 の構成図であり、以下、適応音源符号帳 6 の動作をこの図において説明する。図 3 と同一の部分は同一番号を付す。適応音源符号帳 6 は第一の符号探索手段が選択された場合、第一の量子化音源信号 9 を第一の音源記憶手段 5 1 に格納し、第二の符号探索手段が選択された場合、第二の量子化音源信号 1 0 を第二の音源記憶手段 5 2 に格納しておく。切換手段 5 3 は第一の音源記憶手段 5 1 と第二の音源記憶手段 5 2 に格納されている信号をそれぞれ適応音源信号 7 として出力する。第二の符号探索手段 8 は合成音声と入力音声の歪を小さくする適応音源信号 7 を選択し、選択結果を符号探索結果 1 3 に含めて出力する。

【0035】すなわち第二の符号探索手段 8 は、選択信号 8 a を適応音源符号帳 6 に出力し、第一の音源記憶手段 5 1 と第二の音源記憶手段に記憶された音源信号を切り換える。第一の音源記憶手段 5 1 は、第一の符号探索手段から出力された最新の量子化音源信号 9 を記憶している。第二の音源記憶手段 5 2 は第二の符号探索手段から出力された最新の量子化音源信号を記憶している。第

二の符号探索手段 8 は選択信号 8 a を出力し、符号語切
換手段 5 3 を動作させることにより、第一の音源記憶手
段 5 1 と第二の音源記憶手段 5 2 に記憶されている量子
化音源信号をそれぞれ入力し、両方の量子化音源信号に
基づいて符号化を試みる。その結果、より歪の小さい方
を用いて、符号化を行い符号探索結果として出力する。

【0036】また復号化部の適応音源符号帳 4 0 は図 4
における適応音源符号帳 6 と同一のものである。復号化
部の第二の復号化手段 4 3 は、符号化結果 3 9 に従い適
応音源信号符号帳 4 0 の中から適応音源信号 4 1 を選択
する。

【0037】実施例 4. 上記実施例 3 においては、二つ
の符号探索手段がある場合を示したが、三つ以上の符号
探索手段が存在する場合には、図 4 に示した適用音源符
号帳 6 の内部には音源記憶手段がそれぞれの符号探索手
段に対応して存在し、切換手段 5 3 はこれら 3 つ以上の
音源記憶手段に記憶された量子化音源信号を切り換える。

【0038】実施例 5. 図 5 はこの発明に係わる音声符
号化復号化装置の一実施例における MA 符号帳 2 6 の構
成図であり、以下、MA 符号帳 2 6 の動作をこの図にお
いて説明する。図 1 と同一の部分は同一番号を付す。M
A 符号帳 2 6 は複数の副符号帳 5 4 を持ち、副符号帳切
換手段 5 5 は入力されたピッチ周期 1 5 に応じて副符号
帳の一つを選択し、選択された副符号帳内の MA 符号語
2 7 を出力する。

【0039】例えば図において、副符号帳 1 は 1 5 m s
のピッチ周期に対応する MA 符号帳を格納する。また副
符号帳 2 には 1 6 m s のピッチ周期に対応する符号帳を
格納する。また、副符号帳 3 には 1 7 m s のピッチ周期
に対応する符号帳を格納する。このようにして 1 5 m s
から例えば 2 5 m s までの符号帳を格納しておき、副符
号帳切り換え手段 5 5 は入力されたピッチ周期 1 5 に基
づき、副符号帳を選択できる。例えば、ピッチ周期 1 5
が 1 6 m s である場合には、副符号帳切り換え手段 5 5
は副符号帳 2 を選択しこれを MA 符号語 2 7 として出力
する。なお復号化部の MA 符号帳 3 4 も MA 符号帳 2 6
と同一の構成である。また、AR 符号帳または音源モデ
ル符号帳または駆動音源符号帳を図 5 と同様の構成にす
ることも可能である。

【0040】以上のように、この実施例が特徴とする点
は、ピッチ周期に応じて各符号帳内に複数の符号帳を用
意している点である。音声は例えば、男女の差、あるいは
会話のスピード、あるいは音声の高低等の特徴を有し
ているが、入力される音声は男であるか女であるか、あ
るいは速いか遅いか、あるいは高いか低いかというよう
な特徴はピッチ周期に反映されることが多い。すなわ
ち、ピッチ周期に応じて MA や音源モデルの対応が異な
るという事実がある。このピッチ周期に応じて、複数の
MA 符号帳や音源モデル符号帳を用意しておき、ピッチ

周期に応じた符号帳を用いて量子化するのがこの実施例
の特徴である。

【0041】実施例 6. 図 6、図 7 はこの発明に係わる
音声符号化復号化装置の一実施例の構成図であり、以
下、本発明の動作をこの図において説明する。図 6 は符
号化部を示しており、図 7 は復号化部を示している。図
1、図 2 と同一の部分は同一番号を付し説明を省略す
る。符号化部では、当該フレームが有声音の先頭フレー
ムの場合、音源開始位置 2 0 を当該フレームに先行する
無声フレームにおいて復号化部に伝送する。すなわち、
有音声の先頭フレームを符号化して伝送する前に音源開
始位置 2 0 を符号化して伝送する。このため、有音声の
フレームの伝送が 1 フレーム分遅れることになるが、1
フレーム分の伝送の遅れがあっても復号化部での復号の
時刻がずれるだけであり、復号の品質に影響は少ない。
なお、音源開始位置 2 0 の符号化を有声音の先頭フレー
ムの符号化とともに行ってもよい。ただし、この場合
は、有声音の先頭フレームの符号化情報量が音源開始位
置 2 0 の符号化により減少する。

【0042】復号化部では、第一の復号化手段 3 3 が優
勢音の先頭フレームを復号することが予め定められてお
り、当該フレームが有声音の先頭フレームの場合、第一
の復号化手段 4 3 は入力された音源開始位置 2 0 から第
一の量子化音源信号 3 2 を並べて復号音声 3 7 を生成す
る。

【0043】この例を従来例で示した図 1 7 を用いて説
明する。従来例においては、復号動作はフレームの先頭
から行われていたが、この実施例によれば、音源開始位
置 2 0 から復号がスタートするため、図 1 7 に示した時
刻 T 1 から復号されることになる。従来は時刻 F 1 から
復号されていたため時刻 F 1 から T 1 の間の復号は本来
必要無いにも拘らず、合成音声が生じていたのに対し、
この実施例によれば、音源開始位置 2 0 すなわち時刻 T
1 から復号が開始されるため、本来必要でない部分の合
成音声がなくなる。また、スタート復号開始位置が一致
するため、ピッチ周期のズレもなくなり、復号音声の品
質が向上する。図 1 7 に示したように、入力音声 1 のピ
ッチ周期 P 1、P 2、P 3 は復号音声 3 7 のピッチ周期
Q 1、Q 2、Q 3 とズレているため、このズレが復号音
声の品質を劣化させる原因となっている。これに対し、
この実施例によれば、復号音声も音源開始位置 2 0 より
スタートするためピッチ周期は入力音声のものと同一に
なり、復号音声の品質を向上させるのに役立つ。

【0044】実施例 7. 上記実施例 6 においては、音源
開始位置 2 0 を第一の復号化手段 3 3 に入力する場合を
示しているが、音源開始位置 2 0 を第二の復号化手段 4
3 に入力するようにしてもかまわない。実施例 6 の場合
は前述したように第一の符号化手段が有声音の先頭フレ
ームを複合するということが前提となっているため、音
源開始位置 2 0 を第一の復号化手段 3 3 にのみ入力すれ

ば良かったが、もし、有声音の先頭フレームを、第一と第二の符号化手段のいずれかどちらが符号化するか不明な場合には、音源開始位置 20 を第一と第二の復号化手段の両方に入力させてやることにより、第一と第二の復号化手段のいずれかがその音源開始位置 20 を用いて復号を開始することが可能になる。

【0045】実施例 8. 実施例 1 ないし実施例 7 では、符号化手段選択手段において第一の符号探索手段が選択された場合、適応音源符号帳に第一の量子化音源信号を格納するが、第一の量子化音源信号で MA フィルタを駆動した信号を格納することも可能である。

【0046】実施例 9. 実施例 1 ないし実施例 8 ではスペクトルパラメータとして AR と MA を用いているが、AR のみ、ケプストラム等他のスペクトルパラメータを用いることも可能である。

【0047】実施例 10. 上記実施例 5 においては、図 1 または図 2 に示した音声符号化復号化装置に対して用いられる各符号帳に複数の符号帳を用意する場合を示したが、これら複数の符号帳をピッチ周期で切り換える方式は、図 8 及び図 9 に示した従来の音声符号化復号化装置に対しても用いることができる。すなわち、ピッチ周期 15 に応じて符号帳を選択するという方式は実施例 1 に示した方式あるいは従来の方式に拘らず、他の方式に対しても適用することが可能である。

【0048】実施例 11. 上記実施例においては、音声符号化復号化装置として音声符号化され、かつ復号化される装置の場合について説明したが、単に符号化のみを行う符号化装置、あるいは復号化のみを行う復号化装置においても、それぞれの実施例の符号化部分及び復号化部分を適用することが可能である。

【0049】実施例 12. また上記実施例においては、音声を符号化復号化する場合について説明したが、この発明における音声とは人間が声道から発声する音声に限らず動物や獣等の人間以外の生物が発声する声についても適用することが可能である。同様に生物が発声する声に限らず音として入力されるものであれば、これらの音を符号化復号化する場合にも適用されることが可能である。例えば、楽器の音や摩擦音等の音を入力して符号化、復号化する場合でもかまわない。また、音は人間が知覚出来る場合に限らず人間の耳には感知できない超音波あるいは低音波等の音であってもかまわない。

【0050】評価実験例. 図 8 にサブフレーム数が 2 の場合の本評価実験の構成図を示す。図において MODE 0 は FVQ-GARMA 方式を意味し、MODE 1 は CELP 方式を意味するものとする。まず LSP (AR パラメータ) をフレームに数組予備選択する。次にそれぞれの LSP に対し、SNR seg の良い Mode をサブフレーム毎に選択する。最終的にフレーム全体の SNR seg を最大にする LSP と各サブフレームでの MODE の組み合わせが選択される。

【0051】図 9 に MODE 0 の内部構成図を示す。MODE 0 は FVQ-GARMA と同様で、予備選択された声帯音源波モデルコードと MA コードの全ての組み合わせの中から SNR seg を最大にするものを選択する。

【0052】図 10 に MODE 1 の内部構成図を示す。MODE 1 の CELP は MODE 0 と同期をとるためにピッチ同期の処理とした。まず、適応コードブックの 1 ピッチ長を繰り返しベクトル P とする。なお、前サブフレームが MODE 0 であった場合は声帯音源波モデルで MA フィルタを駆動したものが適応コードブックとなる。次に駆動音源コードブックの 1 ピッチ長を繰り返しベクトル C とする。そして SNR seg を最大にするベクトル P と C のゲインの比を決定する。なお駆動音源コードブックの学習には、短周期予測残差信号を用いた。無声フレームではサブフレーム長が長いサブフレームを複数に分割し、白色雑音で駆動する。なお、適応コードブックは使用しない。コードブックの学習には男女各 5 名が発声した日本語短文 20 文章を用いた。ただし駆動音源コードブックの学習には、10 文章を用いた。評価用データには学習外の日本語短文 10 文章 (学習に用いなかった男女各 5 名が異なる 1 文章ずつを発声) を用いた。

【0053】本方式の性能を調べるために図 11 の条件で合成音を作成した。ただし、音源ピーク位置は自動抽出した。図 12 に本方式 2.4 Kbps (MGARMA) の SNR seg と DC (Cepstrum Distortion) を男女別に示す。この結果をみると男性の方が SNR seg と CD とともに劣っている。これは、ピッチ変動が大きい男性において MODE 1 が多く選択され、MODE 1 においてピッチ長の補間を行わず同一ピッチ長による音源の繰り返しを行っているために音源ピーク位置のずれが大きくなるのが主な原因だと考えられる。

【0054】本方式の合成音声の主観品質を調べるために、図 11 に示す条件で合成音声を作成し、被験者 6 名による簡単な対比較試験を行った。音源ピーク位置を自動抽出した本方式 2.4 Kbps (MGARMA) と音源ピーク位置にマニュアルで修正を加えた従来の 2.4 Kbps FVQ-GARMA (MFVQ)、音源ピーク位置を自動抽出した従来の 2.4 Kbps FVQ-GARMA (AFVQ) 及び 4.8 Kbps CELP 基本方式 (CELP) との比較を行った。試験結果を図 13 に示す。

【0055】本方式は AFVQ より良好であり、音源ピーク位置の抽出誤りに対するロバスト性の向上が確認された。本方式と MFVQ の比較において本方式が選択されない文章では、部分的に残響感が感じられた。音質のばらつきにおいては、MFVQ より今回方式の方が安定していることを確認している。音源ピーク位置抽出や補

間の改良によりMFVQと同等レベルに達すれば4.8 KCELPと同等の品質が得られることが期待できる。

【0056】

【発明の効果】以上説明したように請求項1記載の発明によれば、2種類の符号化方式の中からよりよい方式を選択するので、1種類の符号化方式の場合よりもよりよい符号化が行える。

【0057】また、請求項2記載の発明では符号化手段選択手段により第二の符号探索手段と第一の符号探索手段のうち合成音声と入力音声の歪を小さくする方を選択するため、音源モデルの適合の悪い話者において、復号音声の品質が改善する。

【0058】請求項3記載の発明の適応音源符号帳は適応音源信号として第一の量子化音源信号と音源信号を格納し、第二の符号探索手段が合成音声と入力音声の歪を小さくする方を選択して用いるため、復号音声の品質が改善する。

【0059】また、請求項4、5記載の発明の音声符号化復号化装置では、符号化部と復号化部のスペクトル符号帳、音源モデル符号帳、駆動音源符号帳がピッチ周期に対応して作成された複数の副符号帳を持ち、この副符号帳をピッチ周期により切り換えて用いるため、復号音声の品質が改善する。

【0060】また、請求項6、7記載の発明の音声符号化復号化装置では、有声音の先頭のフレームでの音源開始位置を有声音フレームに先行する無声音フレームで復号化部に出力し、復号化部の第一の復号化手段が第一の量子化音源信号を音源開始位置に同期して復号音声を生成するため、有声音フレームの伝送量を増加させずに符号化部の合成音声と同一の復号音声を生成し、復号音声の品質が改善する。

【図面の簡単な説明】

【図1】この発明の実施例1の音声符号化復号化装置を示す構成図である。

【図2】この発明の実施例1の音声符号化復号化装置を示す構成図である。

【図3】この発明の実施例1の適応音源符号帳を示す構成図である。

【図4】この発明の実施例3の適応音源符号帳を示す構成図である。

【図5】この発明の実施例5のMA符号帳を示す構成図である。

【図6】この発明の実施例6の音声符号化復号化装置を示す構成図である。

【図7】この発明の実施例6の音声符号化復号化装置を示す構成図である。

【図8】この発明に基づく評価実験の音声符号化復号化装置を示す構成図である。

【図9】この発明に基づく評価実験の音声符号化復号化装置を示す構成図である。

【図10】この発明に基づく評価実験の音声符号化復号化装置を示す構成図である。

【図11】この発明に基づく評価実験の条件を示す図である。

【図12】この発明に基づく評価実験のSNRsegとCDを示す図である。

【図13】この発明に基づく評価実験の結果を示す図である。

【図14】従来の音声符号化復号化装置を示す構成図である。

【図15】従来の音声符号化復号化装置を示す構成図である。

【図16】従来の音声符号化復号化装置の符号探索手段の動作を説明する図である。

【図17】従来の音声符号化復号化装置の復号化手段の動作を説明する図である。

【図18】従来の音声符号化復号化装置の問題点を説明する図である。

【符号の説明】

- 1 入力音声
- 2 駆動音源符号帳
- 3 駆動音源符号語
- 4 駆動音源生成手段
- 5 駆動音源信号
- 6 適応音源符号帳
- 7 適応音源信号
- 8 第二の符号探索手段
- 9 第一の量子化音源信号
- 10 第二の量子化音源信号
- 11 AR符号帳
- 12 AR符号語
- 13 第二の符号探索結果
- 14 ピッチ周期抽出手段
- 15 ピッチ周期
- 16 符号化手段選択手段
- 17 符号化手段選択信号
- 18 符号化結果
- 19 音源開始位置抽出手段
- 20 音源開始位置
- 21 音源モデル符号帳
- 22 音源モデル符号語
- 23 音源モデル生成手段
- 24 第一の量子化音源信号
- 25 第一の符号探索手段
- 26 MA符号帳
- 27 MA符号語
- 28 第一の符号探索結果
- 29 音源モデル符号帳
- 30 音源モデル符号語
- 31 音源モデル生成手段

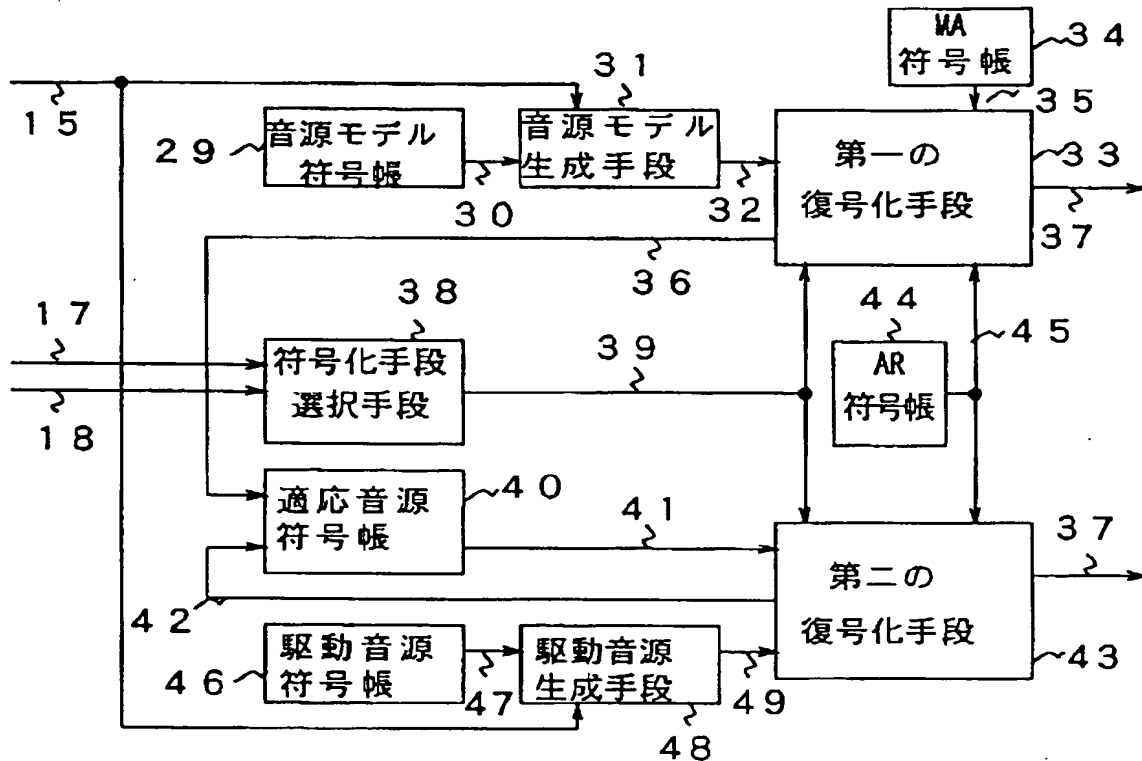
19

20

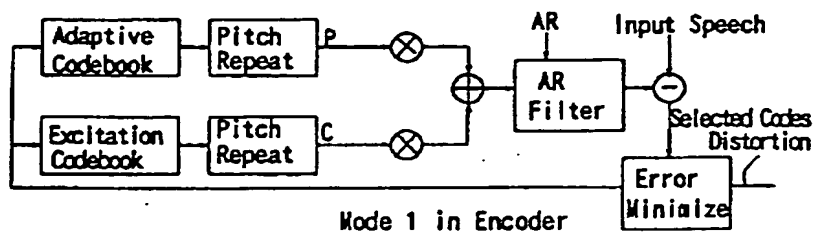
3 2 第一の量子化音源信号
 3 3 第一の復号化手段
 3 4 MA符号帳
 3 5 MA符号語
 3 6 第一の量子化音源信号
 3 7 復号音声
 3 8 復号化手段選択手段
 3 9 符号化結果
 4 0 適応音源符号帳
 4 1 適応音源信号
 4 2 第二の量子化音源信号
 4 3 第二の復号化手段
 4 4 AR符号帳
 4 5 AR符号語
 4 6 駆動音源符号帳

4 7 駆動音源符号語
 4 8 駆動音源生成手段
 4 9 駆動音源信号
 5 0 音源記憶手段
 5 1 第一の音源記憶手段
 5 2 第二の音源記憶手段
 5 3 切換手段
 5 4 副符号帳
 5 5 副符号帳切換手段
 5 6 符号化結果
 5 7 符号探索手段
 5 8 復号化手段
 5 9 量子化音源信号
 6 0 量子化音源信号

【図2】



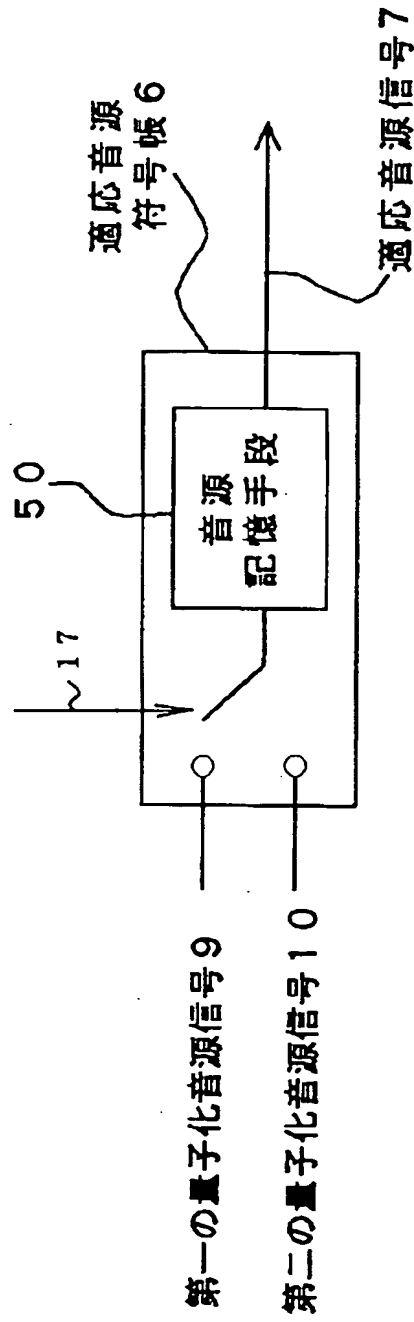
【図10】



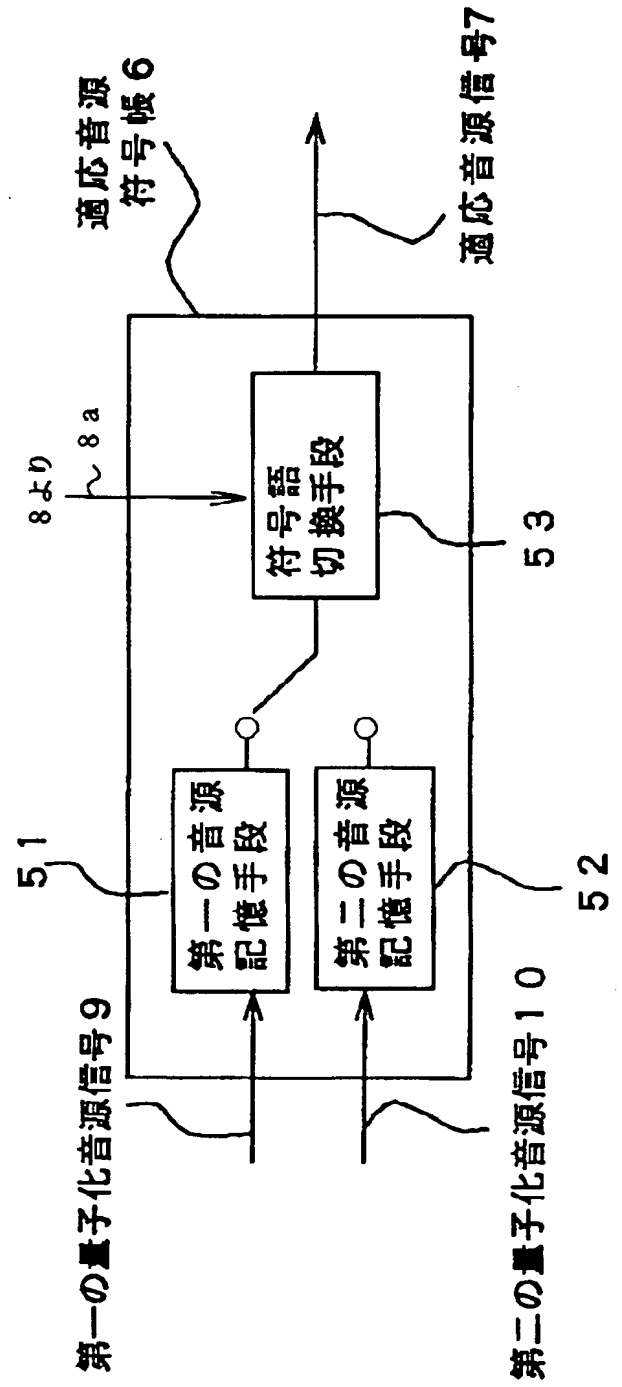
【図12】

SNRseg and CD		
	SNRseg[dB]	CD [dB]
MALE	7.8	3.69
FEMALE	10.0	3.45

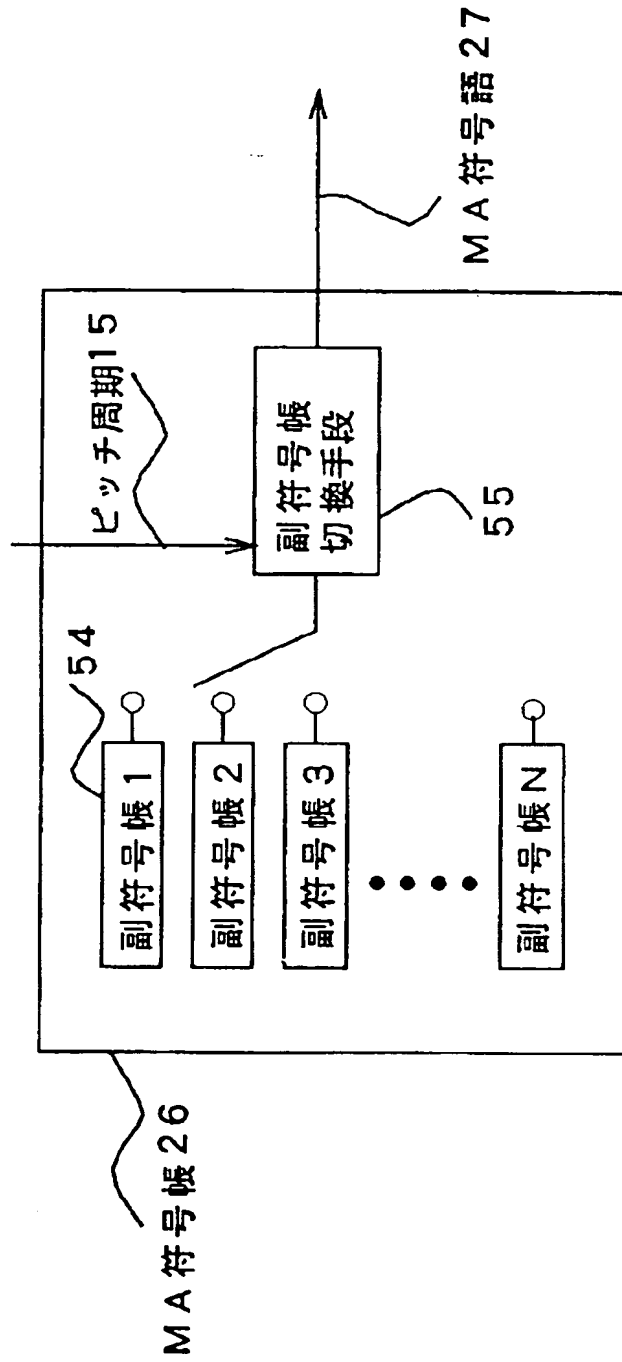
【図3】



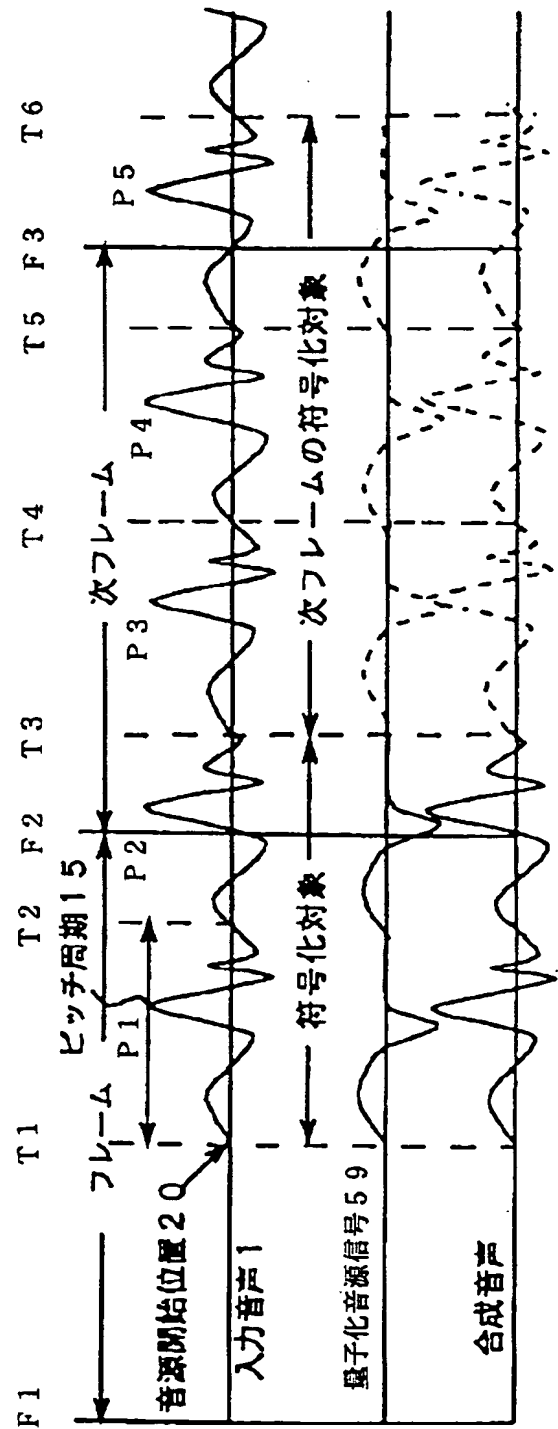
【図4】



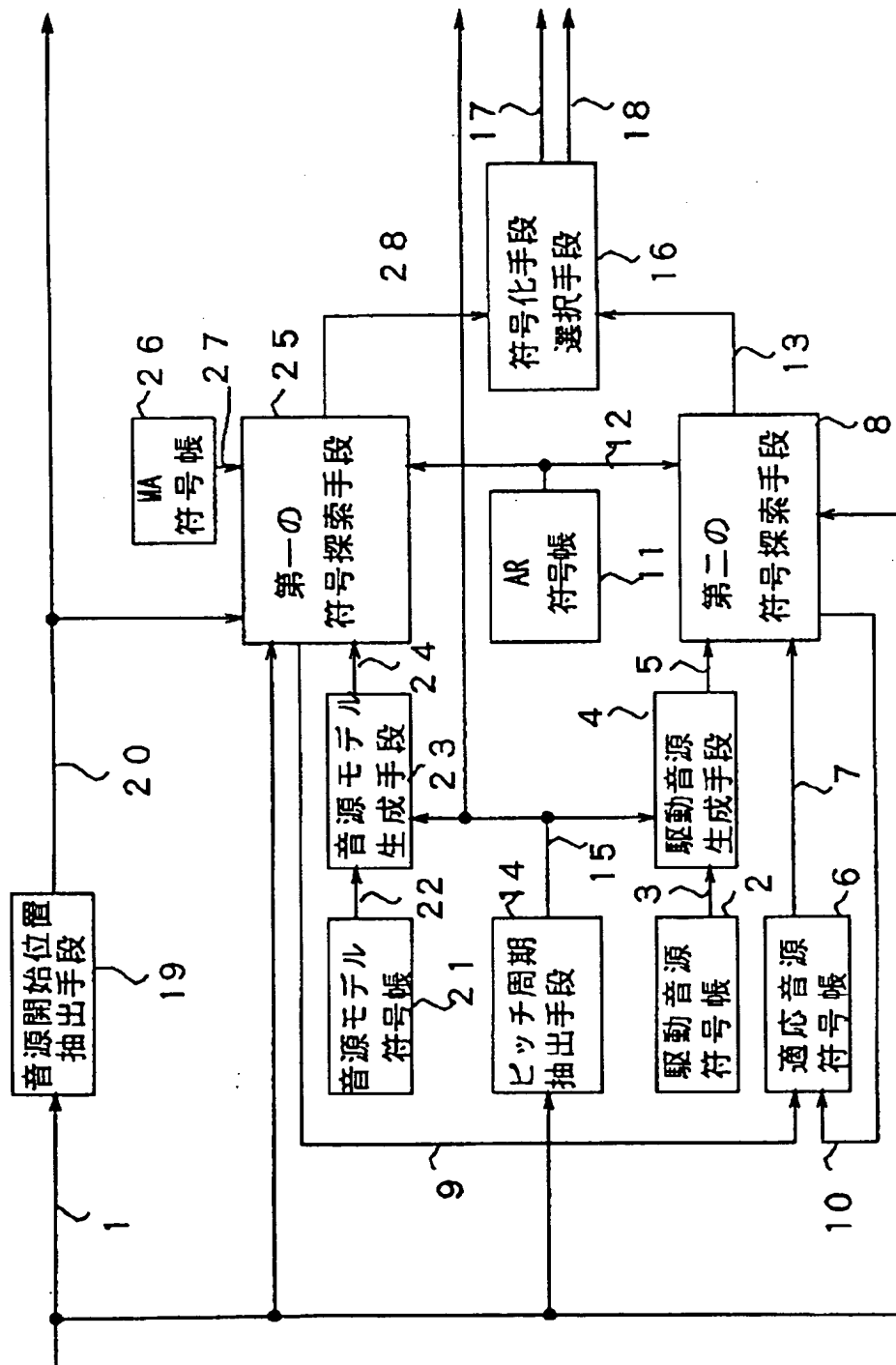
【図 5】



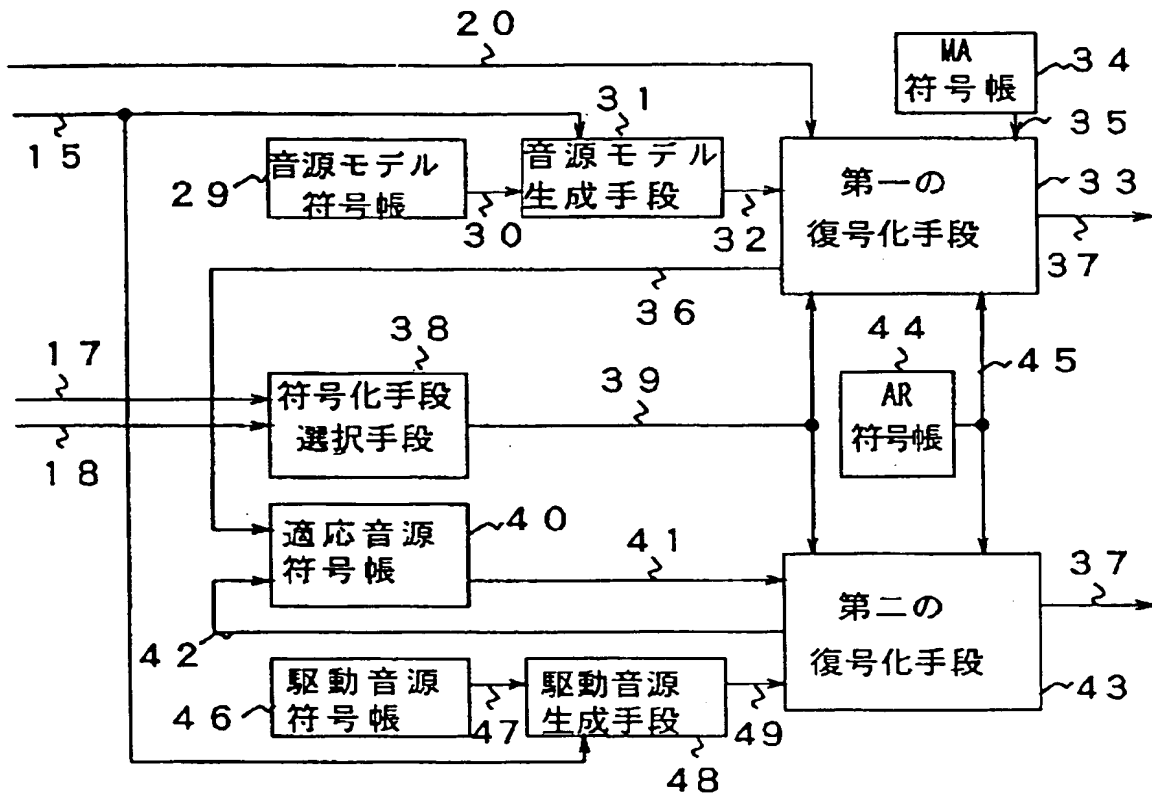
【図 16】



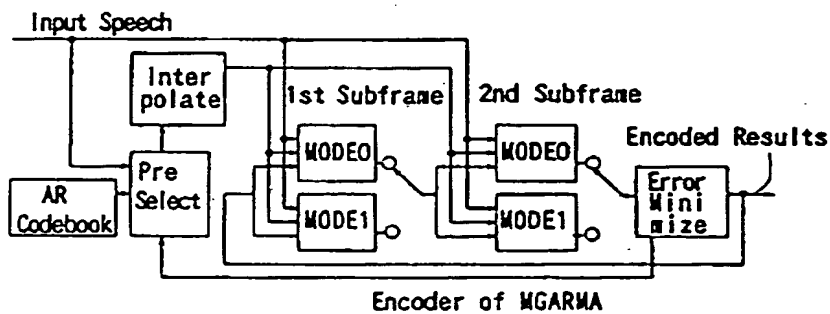
【図 6】



【図7】



【図8】

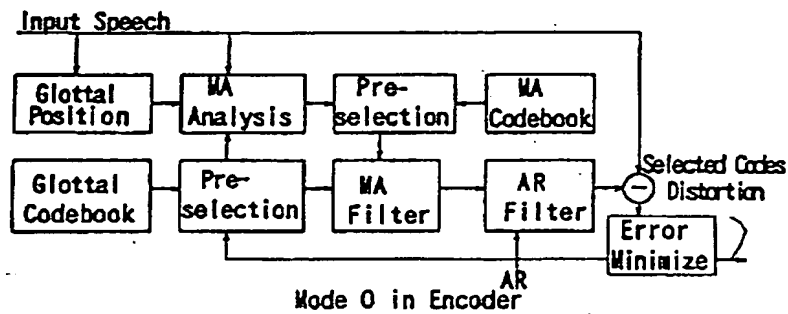


【図13】

Preference Score [%]

MGARMA 2.4 vs AFVQ 2.4	82 : 18
MGARMA 2.4 vs MFVQ 2.4	30 : 70
MGARMA 2.4 vs CELP 4.8	30 : 70

【図9】

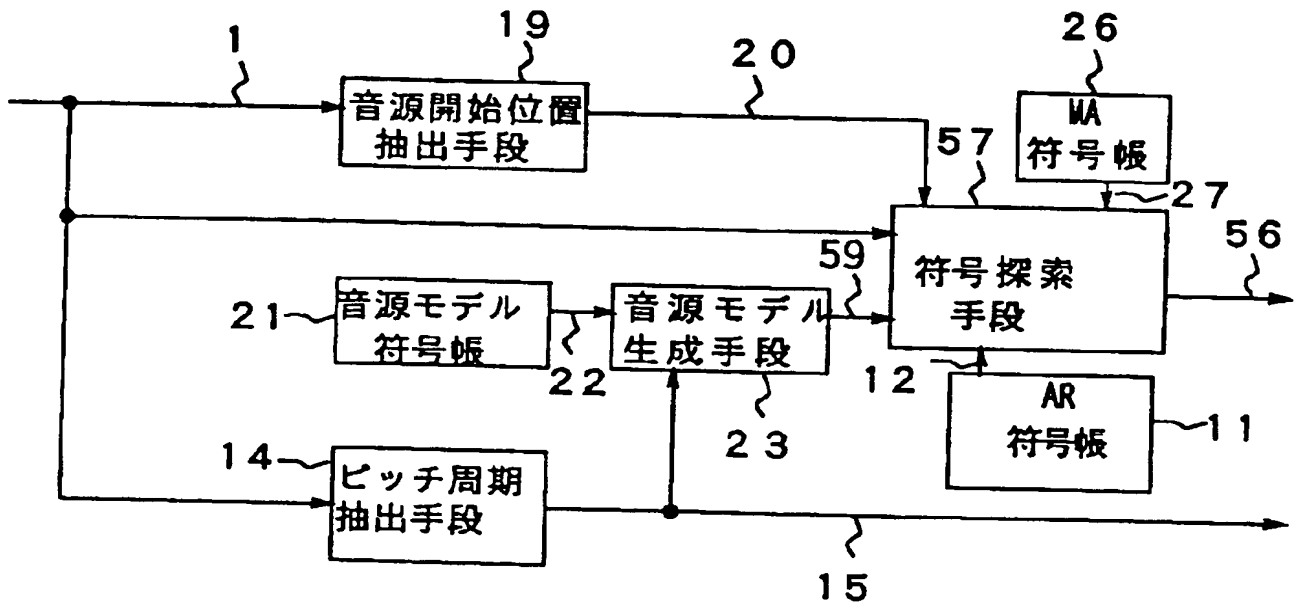


【図11】

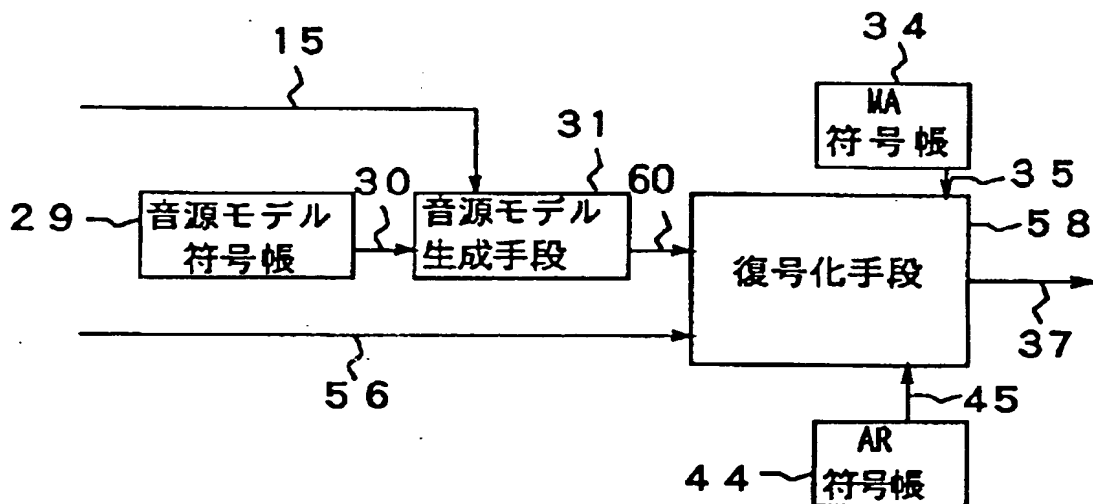
Simulation Condition

Sampling Frequency	8 KHz
Bit Rate	2.4 Kbps
Frame Length	37.5 ms
Subframe Length	12.5 ms
Pitch & Voice/Unvoice	12 bit
LPC Analysis Order	10
LSP Code (ADVQ(4)) / frame	11 bit
Number of Pre-selection	4
Voice	
Power / subframe	4 bit
Flag of Mode / subframe	1 bit
Mode 0 (FVQ-GARMA)	
MA Code / subframe	9 bit
Number of Pre-selection	64
Glottal Code / subframe	8 bit
Number of Pre-selection	16
Mode 1 (CELP)	
Ratio of Adaptive Gain to Excitation Gain / subframe	3 bit
Excitation Code / subframe	14 bit
Unvoice	
Pulse Position for GARMA	7 bit
Excitation Gain / Subframe	4 bit
Excitation Code / Subframe	10 bit

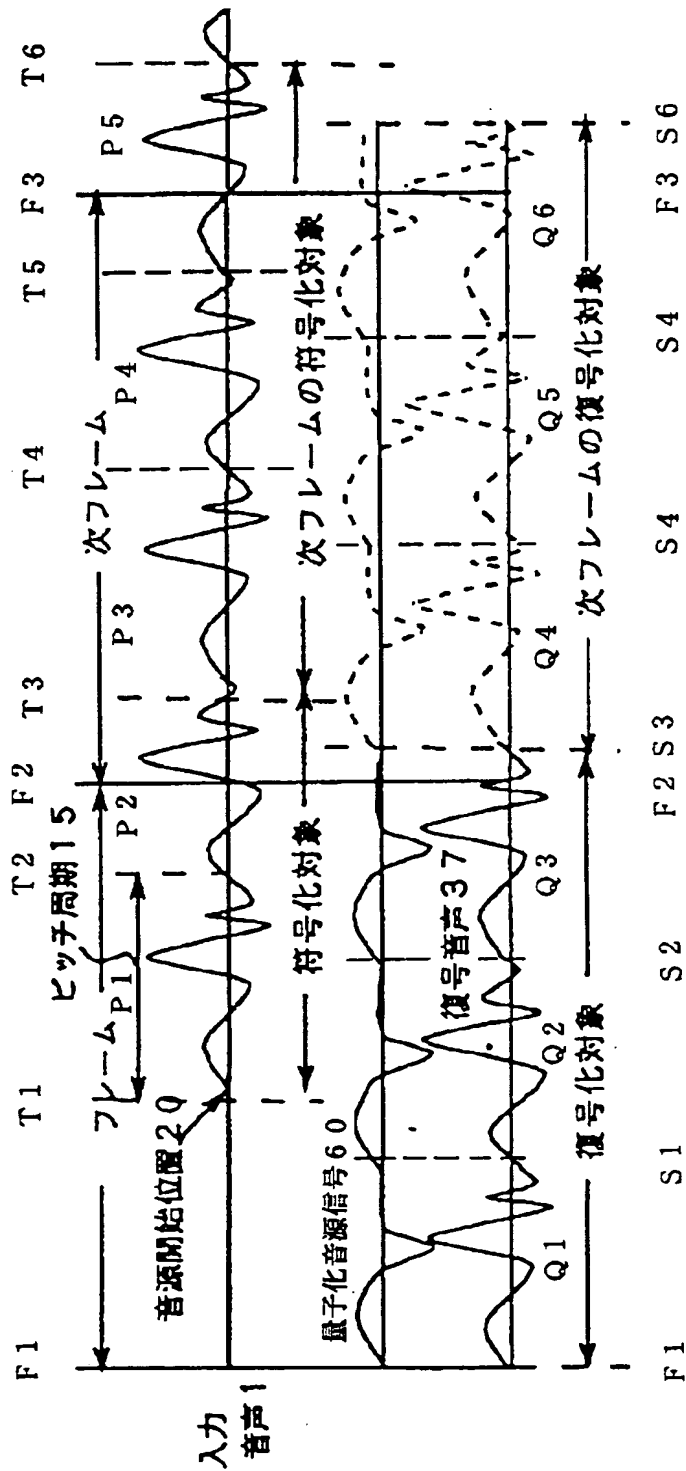
【図 1 4】



【図 1 5】



【図 17】



【図 18】

